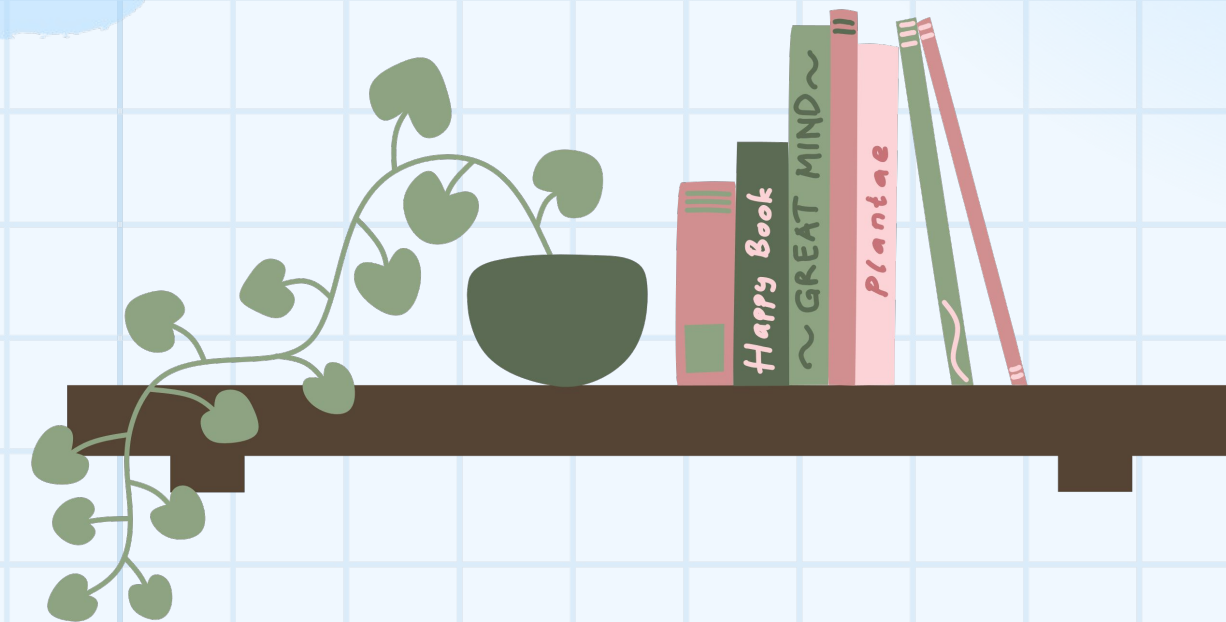
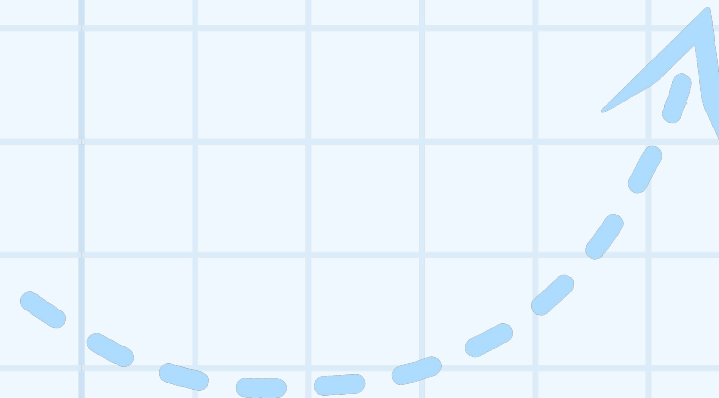
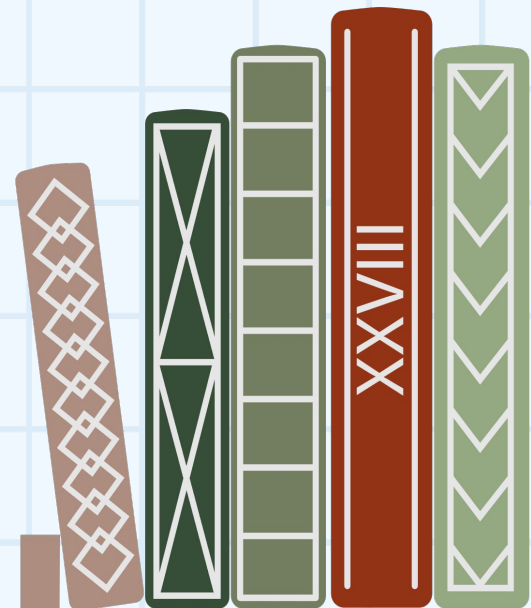




# BS./BSC.IN

Applied ai and Data Science

## Basics of Data Analytics



# Let's dive into and learn:



1

## Errors or Inconsistencies



# Data errors and inconsistencies

- After removing duplicates, formatting data and handling missing values, we have to look for errors and inconsistencies in the data
- More difficult to do
- Not always obvious which ones are errors
- Need to use our common sense and intuition to check

# Data Errors



- How would you know if some data was obviously wrong?
- Need to know the context and source of collection
- Check for the same information from another source

# Data Errors



Name	Age	Salary	Location
Priya	29	55000	Patna
Rahul	32	67000	Mohali
Devesh	28	52000	Bhubaneswar
Bhavna	-31	64000	Thrissur
Priya	29	55000	Patna

# Data Errors



Name	Age	Salary	Location
Priya	29	55000	Patna
Rahul	132	67000	Mohali
Devesh	28	52000	Bhubaneswar
Bhavna	31	64000	Thrissur
Priya	29	55000	Patna



# Data Discrepancies

Database 1

ID	Name	Age	Height
GH89765	Priya	29	63
RF23876	Rahul	32	61
UT75097	Devesh	28	63
FC37694	Bhavna	31	59

Database 2

ID	Name	Age	Height
GH89765	Priya	29	63
RF23876	Rahul	32	61
UT75097	Devesh	28	68
FC37694	Bhavna	31	59



Powered by



FutureSense

# Data Errors and Discrepancies

- Go back to the source to figure out which one is correct
- Fix any errors you see
- May need to drop records if you can't figure out the error or discrepancy

# Recap



Spotting data errors or  
discrepancies



# Thank you

